*Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference*
*Edited by* Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
©2025 ESREL SRA-E 2025 Organizers. *Published by* Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P1228-cd

# Intelligent Anomaly Detection for Drivetrain Systems in Wind Turbines

Zifei Xu

*School of Engineering, University of Liverpool, UK. E-mail: zifei.xu@liverpool.ac.uk*

He Li

*School of Engineering, Liverpool John Moores University, UK. E-mail: he.li@ljmu.ac.uk*

Jin Wang

*School of Engineering, Liverpool John Moores University, UK. E-mail: j.wang@ljmu.ac.uk*

Musa Bashir *

*School of Engineering, University of Liverpool, UK. E-mail: m.b.bashir@liverpool.ac.uk*

**Abstract:** The safety and reliability of the drivetrain system in offshore wind turbines are crucial for their effective operation. Detecting anomalous behaviour within the drivetrain and providing reliable prognostic information can significantly reduce the risk of severe failures. Ensuring the reliability and safety of intelligent models is of paramount importance in the AI-driven, data-centric era. To address this challenge, this paper presents an intelligent anomaly detection model capable of issuing alerts prior to abnormal shutdowns, thereby ensuring system safety. A physics-informed probabilistic neural network was developed, integrating physical insights into the neural framework to manage prediction uncertainty and enhance the safety and reliability of failure alarms generated by the intelligent model. Overall, the proposed method offers a more reliable prognostic framework to enhance the safety and stability of wind turbines, including offshore installations, during operation while reducing costs

***Keywords***: Reliable prognostics, anomaly detection; physics-informed neural network, offshore wind turbine, drivetrain, SCADA data

**Nomenclature**

| | |
|---|---|
| Bi-LSTM | Bidirectional Long Short-Term Memory |
| CNN | Convolutional Neural Network |
| CLSTM | Convolutional Long Short-Term Memory |
| GRU | Gated Recurrent Unit |
| LSTM | Long Short-Term Memory |
| Py-PINet | physics informed probabilistic Informer network |
| RFR | Random Forest Regression |
| SVR | Support Vector Regression |
| TransNet | Transformer Network |
| Py-PINet | physics informed probabilistic Informer network |

## 1. Introduction

Offshore wind plays a pivotal role in the clean energy transition, particularly in Europe. In 2023, Europe reached a milestone, installing 3.8 GW of new offshore wind capacity across six markets, bringing the region's total capacity to 34 GW(Arefin and Ishraque 2023). The UK and Germany remain at the forefront, contributing 43% and 24% of Europe's offshore wind capacity, respectively. As Europe pursues ambitious renewable energy targets, offshore wind is increasingly critical in reducing reliance on fossil fuels and enhancing energy security, solidifying the region's leadership in clean energy development.

The drivetrain system, comprising the rotor, gearbox, and generator, is the core operational system in wind turbines (WTs), with its reliability directly influencing the WTs' safe, stable, and efficient operation(Mehlan and Nejad 2023). For offshore WTs, the drivetrain has even higher reliability requirements due to the more challenging operational environment(Barter et al. 2023). Therefore, it is crucial to study and develop effective and reliable prognostics and health management approaches to ensure the health of the drivetrain system.

Anomaly detection, as a foundational task for fault diagnosis, is crucial for assessing the health status of equipment, especially in scenarios where predefined fault patterns or labeled data are unavailable. In the era of data-driven technologies, anomaly detection has gained significant attention and importance for wind turbines, particularly due to its ability to utilize diverse and large-scale data streams provided by Supervisory Control and Data Acquisition (SCADA) systems, such as temperature, vibration, and power output. The development and application of effective anomaly detection methods (Zhang et al. 2018; Campoverde-Vilela et al. 2023; Shanbr et al. 2018; Moghadam and Nejad 2022), which have greatly enhanced the reliability and safety of wind turbine drivetrains, ensuring their stable operation under varying conditions.

Compared to the aforementioned traditional methods, deep learning-based intelligent anomaly detection methods offer powerful nonlinear modeling capabilities and can effectively identify complex temporal and spatial patterns(Z. Xu et al. 2022; Zifei Xu et al. 2024). As a result, they are better suited for handling anomaly detection in offshore WTs' drivetrains(Yan, Liu, and Ren 2023; Xiang et al. 2021; Z. Y. Zhang and Wang 2014), which have achieved robust fault detection for wind turbine drivetrain mechanical systems.

However, these studies focus primarily on adopting advanced techniques from fields like image processing and natural language processing to improve the performance of intelligent models in anomaly detection tasks. However, they often overlook the fact that the monitored equipment is a physical entity, where the physical relationships between monitored variables and predicted outcomes should be considered. Furthermore, these AI-driven methods do not address how the inherent reliability of the AI models themselves may influence the trustworthiness of the anomaly detection results.

Therefore, in this study, a fully intelligent solution for anomaly detection applied to the drivetrain system of wind turbines (WTs) is proposed, leveraging a physics-informed probabilistic neural network. The main contributions are as follows:

1. A physics-informed neural network model is proposed, where physical information is embedded into the network through the diffusion model of spatiotemporal positions from virtual sensors, enhancing predictions for the target physical sensors.
2. A physics-informed probabilistic neural network model is developed, utilizing Monte Carlo dropout. This model provides confident evaluations for healthy data and generates significant deviations when evaluating anomalous data, offering robust warning signals.
3. A reliable intelligent anomaly detection system is established, enabling effective prognosis for the drivetrain system of wind turbines.

## 2. Methodologies

### 2.1. *Problem definition*

Condition monitoring of the drivetrain using SCADA data typically involves multiple sensors, which are positioned around the drivetrain to capture variations in its operational state. If several sensors exhibit high correlations, their measurements are expected to share a nonlinear mapping relationship. Theoretically, it is possible to predict the measurements of one sensor based on data from others. To facilitate anomaly detection, the concept of a Normal Behavior Model (NBM) is introduced. This model is constructed using healthy data and leverages data from several sensors to build a predictive model. The task is formulated as a regression problem, where the model predicts the measurement of one or more target sensors based on the input from others. The anomaly detection capability relies on the assumption that when the system deviates from its normal state, the output of the NBM will diverge from the observed values. If the prediction error exceeds a predefined threshold (representing the upper limit of acceptable error under healthy conditions), the system is flagged as anomalous.

The mathematical process of the Normal Behavior Model (NBM) is defined as follows: Let the monitored physical sensor data be $x \in \mathcal{R}^{n\times m}$, which is a function of time and spatial state:

$$x = x(t,s), t \in \mathcal{R}^n, s \in \mathcal{R}^m \tag{1}$$

where $t$ represents time,$s$ denotes the spatial coordinates, and $m$ is the spatial dimension, equal to the number of sensors.

Assume the NBM model $\mathcal{N}(\cdot)$ with learnable parameters $\theta$ . The model is optimized through supervised learning by maximum likelihood estimation $\hat{\theta} = \text{argmax}x_{\theta}P(\theta|x)$.

Given $x_{healthy}, x_{healthy} \in \mathcal{R}^{n \times m}$, the error distribution $\varepsilon \in \mathcal{R}^n$ for normal data is estimate as:

$$\varepsilon_{\text{health}} = \text{simlarity}(\hat{y}_{\text{healthy}}, y_{\text{healthy}}) \quad (2)$$

where $\text{simlarity}(\cdot)$ is a distant estimation function. $y_{\text{healthy}}, y_{\text{healthy}} \in \mathcal{R}^{n \times k}$ represents the real target sensor measurement, and $\hat{y}_{\text{healthy}}$ is the estimate of target sensor measurement:

$$\hat{y}_{\text{healthy}} = \mathcal{N}\left(x_{healthy} \middle| \hat{\theta}\right) \quad (3)$$

In this study, the number of target sensors is $k = 1$. The threshold $\tau$ is determined by the upper limit of the reconstruction error for healthy data based on the NBMs.

During the anomaly detection, given $x_{\text{test}} \in \mathcal{R}^{n \times m}$, the prediction by the NBM model is:

$$\hat{y}_{\text{test}} = \mathcal{N}\left(\mathrm{x}_{\text{test}} \middle| \hat{\theta}\right) \quad (4)$$

The error for the monitoring condition is $\varepsilon_{\text{test}}$ using the same equation as Eq. (2). An anomaly is detected when any $\varepsilon_{\text{test}}$ exceeds the threshold.

However, Current AI-driven NBM (Normal Behavior Model) approaches often overlook the temporal and spatial relationships of features, which significantly impact the accuracy of predictions. Moreover, the nonlinear AI-driven models inherently lack physical interpretability, limiting their generalization and predictive performance. Deterministic models are unable to quantify prediction uncertainty, making it difficult to assess the reliability of their outputs effectively. To address these issues, Sections 2.2 and 2.3 propose: (1) a physics-informed probabilistic neural network model and (2) a reliable anomaly detection framework built upon this model.

### 2.2. *Physics Informed Probabilistic Informer Network*

The physics informed probabilistic Informer network (Py-PINet) framework is based on encoder-decoder framework(Zhou et al. 2020). The general framework of the Py-PINet is shown in Figure 1.

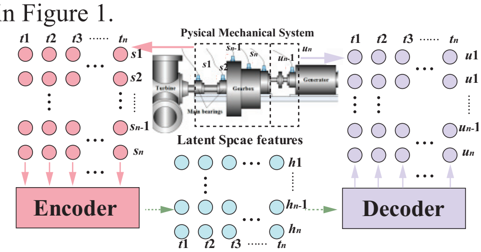



Figure 1 Py-PINet framework

The whole Py-PINet $\mathcal{N}(\cdot)$ consists of an encoder $f(\cdot)$ and a decoder $g(\cdot)$. The encoder $f(\cdot)$ is structured based on the Informer encoder. The hidden feature in the latent space is:

$$h = f(x_{raw} | \theta) \quad (5)$$

where $h \in \mathcal{R}^{n \times p}$, $k$ is the hidden dimensions in the latent space, $x_{raw} \in \mathcal{R}^{n \times m}$. Consider $h(t, x)$ as a function of both time and spatial variables, representing a virtual mechanical space with $p$ virtual sensors $p \in \mathcal{R}^p$, each evolving with $t$. These virtual sensors are designed to predict the temperature variables of a real physical mechanical system.

Like traditional approaches, $h$ is mapped to the temperature $u$ through a decoder $g(\cdot)$, which is used to predict the temperature of the target sensor in the physical mechanical system.

$$\hat{u} = g(h | \theta) \quad (6)$$

where $u \in \mathcal{R}^{n \times k}$, $k$ is number of the target sensors, in this study, the number of target sensors is $k = 1$.

The parameters of the encoder $f(\cdot)$ and the decoder $g(\cdot)$ in the Py-PInet $\mathcal{N}(\cdot)$ can be learned by conventional supervised learning based on data loss $\mathcal{L}_{\text{data}} = \|u - \hat{u}\|^2$.
where $u$ is the real temperature of the target sensor.

In addition to the data loss, a novel physical constraint loss is proposed for Py-PINet training. It is assumed that the temperatures of the target sensors should be approximately following a 1-D diffusion equation, the PDE loss should be:

$$\mathcal{L}_{\text{pde}} = \left\| \frac{\partial u}{\partial t} - \alpha \frac{\partial^2 u}{\partial x^2} \right\|^2 \quad (7)$$

where $\frac{\partial u}{\partial t}$ and $\frac{\partial^2 u}{\partial x^2}$ can be obtained by the chain rule.

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial h} \cdot \frac{\partial h}{\partial t}$$
$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial h^2} \cdot \left(\frac{\partial h}{\partial x}\right)^2 + \frac{\partial u}{\partial h} \cdot \frac{\partial^2 h}{\partial x^2} \quad (8)$$

where $\alpha$ can be approximated by a neural network as well. $\frac{\partial u}{\partial h}$ and its high order differential item can be obtained by auto grad. $\frac{\partial^2 h}{\partial x^2}$, $\frac{\partial h}{\partial x}$ and $\frac{\partial h}{\partial t}$ can be obtained by finite difference.
Therefore, the total loss for training Py-PINet is in Equation (9), where each item has the same contribution.

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{pde}} \tag{9}$$

### 2.3. *Reliable anomaly detection framework*

The proposed reliable anomaly detection framework is developed based on the proposed Py-PINet and uncertainty quantification. MC dropout is utilized as the method for implementing uncertainty prediction within the Py-PINet model. By applying MC dropout, the model can perform uncertainty prediction for each monitoring data. This allows the model to quantify the uncertainty associated with its predictions, thereby enhancing the reliability of temperature prediction corresponding to the target sensors. The quantified uncertainty is used to assess the reliability of the AI-driven model in prognosis, ensuring robust and reliable monitoring of the system.

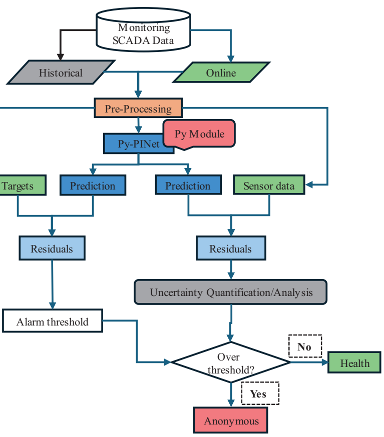



Figure 2 Reliable anomaly detection framework

Figure 2 describes the steps of intelligent anomaly detection. Step 1: SCADA data collection: The process begins with monitoring data from the SCADA system, which consists of both historical data and real-time online data. Step 2: Pre-Processing, historical and online data undergo pre-processing to remove erroneous data and normalize the data, making it suitable for machine learning. Step 3: Model Training (Historical Data): Py-PINet is the proposed model used in this step. During training, it incorporates both data loss and physical PDE loss to ensure accurate predictions. The model learns to predict target values and compute residuals using historical healthy-condition data. Step 4: Residual Analysis: Residuals calculated from historical data are used to define an Alarm Threshold, which serves as the basis for anomaly detection. Step 5: Prediction (Online Data): The trained Py-PINet model processes real-time online data, generating predictions and computing residuals as health indicators, Step 6: Uncertainty Quantification:

The model quantifies uncertainty to provide interval predictions for health conditions, improving a layer of reliability to the detection process. Step 7: Anomaly Detection, The health indicators are compared against the predefined alarm threshold. If the residual exceeds the threshold, it is classified as an anomaly; otherwise, the system is deemed healthy.

## 3. Case Study and Analysis

To validate and analyze the proposed method's effectiveness in anomaly detection for the drivetrain system, this section utilizes three datasets derived from SCADA data recorded during WTs' operation at 1-minute intervals. Boolean variables from the SCADA data are employed to assess whether the wind turbine is functioning under normal conditions. The computational setup comprises a 12th Gen Intel i9-12900K processor and an NVIDIA RTX A5500 graphics card. All AI-driven models were developed using PyTorch version 2.5.1.

During training, the optimizer used was ADAM, with a dropout rate of 0.1. The hidden dimensions in the AI-driven models are 256, with a batch size of 20. The maximum number of training epochs is 300.

### 3.1. *Wind Turbine Dataset*

Two datasets are used in this study to validate the effectiveness and reliability of the proposed method. One dataset contains health data, which is selected from normal operation data of a wind turbine in China, collected between March and August 2019, comprising 195,773 valid samples. This dataset is used to examine the effectiveness and reliability of the proposed model in establishing Normal Behavior Models (NBMs). The other dataset includes abnormal data, which contains records where the wind turbine was alarmed due to a serious abnormality in the drivetrain system, leading to a shutdown for inspection. The original dataset consists of data from 107,321 time points, and after preprocessing, 105,501 valid samples were obtained.

### 3.2. *Normal behaviour performance evaluation*

The primary function of the NBMs is to establish nonlinear relationships among multiple sensors, including the Main Shaft Front Bearing Temperature (MSFBT), Main Shaft Rear Bearing Temperature (MSRBT), Environmental Temperature (ET), Nacelle Temperature (NT), and Hub Temperature (HT). A temperature prediction model was developed to forecast changes in MSFBT based on the other four monitoring variables. The performance of the NBMs is evaluated using Root Mean Square Error (RMSE) and R-squared ($R^2$) metrics

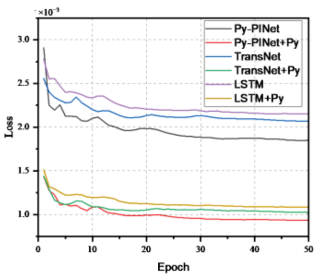



Figure 3 Validation loss

The validation loss curves in Figure 3 demonstrate the effectiveness of the proposed Py (physics-informed enhancement) module. Across all baseline models (Py-PINet, TransNet, and LSTM), the integration of the physics-informed module (+Py) significantly reduces the validation loss, indicating its universal benefits. Notably, Py-PINet achieves the lowest validation loss among all models, outperforming its base version Py-PINet, as well as TransNet(+Py) and LSTM(+Py), highlighting the superiority of the Py-PINet architecture and the physics-informed module. Furthermore, the physics-informed models converge faster and maintain lower losses throughout training, showcasing their efficiency and robustness. These results validate the effectiveness of the proposed physics-informed module in enhancing model performance and the overall reliability of Py-PINet as the best-performing framework.

Table 1 Prediction performance of the NBMs

| NBMs/PEIs | RMSE | R2(%) |
|---|---|---|
| SVR | 1.656 | 85.69 |
| RFR | 1.841 | 83.72 |
| CNN | 1.479 | 90.57 |
| CLSTM | 1.241 | 97.36 |
| GRU | 1.189 | 97.50 |
| Bi-LSTM | 1.170 | 97.58 |
| LSTM(+Py) | 1.195 (1.188) | 97.46(97.49) |
| TransNet(+Py) | 1.178(1.152) | 97.54(97.62) |
| **Py-PINet** | **0.992** | **98.16** |

The experimental results in Table 1 emphasize the effectiveness of the proposed Physics-Informed module (+Py) in enhancing the performance of time series forecasting models. Traditional models such as SVR and RFR show significantly poorer results, with high RMSE values (e.g., 1.656 for SVR and 1.841 for RFR) and low $R^2$ scores (e.g., 85.69% for SVR and 83.72% for RFR), highlighting their limitations for complex prediction tasks. By integrating the +Py module into advanced deep learning models like LSTM and TransNet, substantial performance improvements are observed. For instance, LSTM(+Py) reduces RMSE from 1.195 to 1.188 and increases $R^2$ from 97.46% to 97.49%, while TransNet (+Py) lowers RMSE from 1.178 to 1.152 and improves $R^2$ from 97.54% to 97.62%. These results demonstrate the consistent benefits of incorporating physics-informed enhancements across different architectures. Among all methods, the proposed Py-PINet achieves the best performance, with an RMSE of 0.992 and an $R^2$ of 98.16%, significantly outperforming both traditional models and other deep learning approaches enhanced with the physics-informed module. These findings underscore the strong potential of the physics-informed module in leveraging domain-specific knowledge to reduce uncertainty and enhance model accuracy, making it particularly suitable for complex system forecasting tasks.

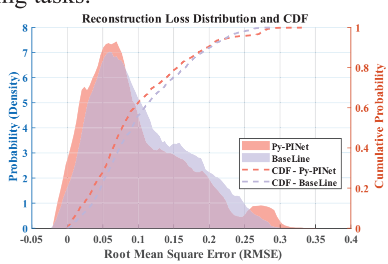
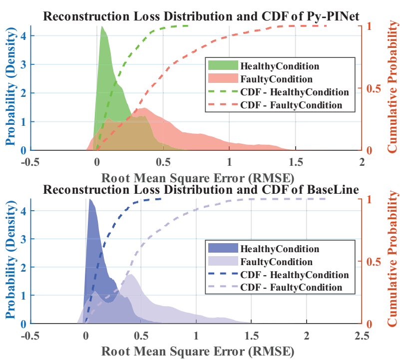



Figure 4 Uncertainty analysis

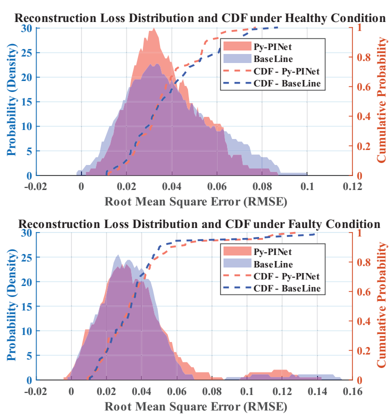
Figure 4 compares the uncertainty of Py-PINet and Baseline (Py-PINet without physics-informed), highlighting the superior performance of Py-PINet in reconstruction accuracy and robustness. From the probability density distribution, Py-PINet exhibits a sharper peak concentrated within the low RMSE range (0–0.1), indicating a higher proportion of samples with minimal prediction errors and fewer extreme outliers. In contrast, Baseline's distribution is more spread out, with a notable presence of higher RMSE values, reflecting less accurate predictions. The cumulative distribution function (CDF) further reinforces this observation, as Py-PINet demonstrates a faster rise in cumulative probability at lower RMSE thresholds, achieving near-total coverage by RMSE = 0.15. This suggests that Py-PINet not only reduces overall prediction errors but also maintains greater consistency across samples. The improvements can be attributed to the enhanced feature extraction capability introduced by the physics-informed module, which likely mitigates noise and captures relevant information more effectively. These results validate Py-PINet as a more reliable and precise approach for NBMs tasks.

### 3.3. *Uncertainty Analysis*

Uncertainty quantification for the prediction of an AI-driven model is critical, which can help to assess its reliability and safety while applying to fault prognosis or anomaly detection. In the rest of the discussion, the baseline means the proposal model without physics-informed module.




(a) Aleatoric uncertainty analysis




(b) Epistemic uncertainty analysis

Figure 5 Uncertainty analysis for healthy and faulty condition

As shown in Figure 5(a), the comparison between Py-PINet and baseline highlights the superior performance of Py-PINet in anomaly detection tasks. As observed from the reconstruction loss distributions, Py-PINet demonstrates a highly concentrated error distribution for healthy conditions, with most RMSE values close to zero, indicating excellent stability and minimal prediction errors. In contrast, the Baseline model shows a wider error distribution for healthy conditions, suggesting lower consistency in its predictions. For faulty conditions, Py-PINet effectively shifts the reconstruction loss distribution to higher RMSE values, achieving a clear separation from healthy conditions. Moreover, Py-PINet's CDF curve increases rapidly for healthy samples, reaching stability at an RMSE of approximately 0.5, while its faulty samples exhibit a slower increase, achieving full coverage at an RMSE of approximately 1.0. This indicates a distinct boundary between healthy and faulty conditions, with minimal overlap. In contrast, the Baseline model exhibits slower CDF growth for healthy samples and significant overlap between healthy and faulty distributions, particularly in the RMSE range below 0.5. This overlap indicates a higher likelihood of misclassification, reducing the model's reliability in differentiating between the two conditions. Additionally, Py-PINet's focused error distribution for healthy conditions and well-separated faulty condition distribution suggest enhanced robustness and reliability. The clear separation between the two conditions ensures lower false positives and negatives, which are critical for effective anomaly detection. In contrast, the Baseline's wider distributions and overlapping CDFs reflect its weaker ability to distinguish between the two states, potentially leading to less accurate detection. These findings underscore the effectiveness of Py-PINet as a more precise and robust solution for anomaly detection tasks.

As shown in Figure 5(b), the comparison between Py-PINet and Baseline demonstrates a clear distinction in model uncertainty under healthy and faulty conditions, aligning well with the characteristics of an effective anomaly detection framework. When measuring healthy data, Py-PINet exhibits significantly lower model uncertainty, as evidenced by its sharp, left-skewed RMSE distribution with a higher peak and smaller spread. Furthermore, the CDF of Py-PINet increases rapidly and stabilizes at lower RMSE thresholds compared to the baseline, indicating that the model is highly confident in reconstructing data within its learned knowledge domain. This aligns with the expected behaviour, where model uncertainty is minimized for in-distribution samples. In contrast, when processing faulty data, Py-PINet reveals higher uncertainty compared to baseline, with a broader RMSE distribution and a slower CDF rise. This behaviour reflects the model's lack of confidence in handling out-of-distribution samples, which serves as a critical warning signal for anomaly detection. The increased uncertainty under faulty conditions highlights Py-PINet's sensitivity to data outside its knowledge domain, thereby offering a valuable mechanism to flag potential anomalies. In summary, Py-PINet effectively balances low uncertainty for healthy data and heightened uncertainty for faulty data, making it a more robust and interpretable choice for anomaly detection compared to baseline.

### 3.4. *Prognosis verification*

This section examines the prognosis capability of the proposal method in real wind turbine anomaly detection.

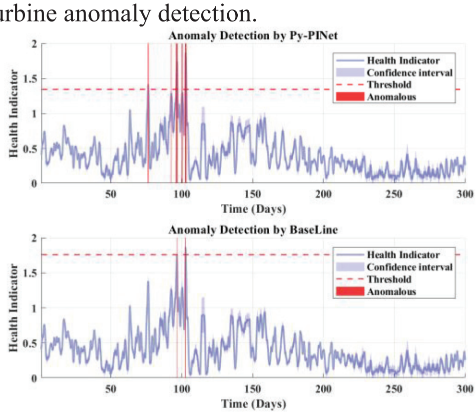



Figure 6 Anomaly detection

Figure 6 makes a comparison between Py-PINet and Baseline highlights significant differences in their anomaly detection capabilities, particularly in terms of missed anomalies and prognosis functionality. Py-PINet demonstrates superior performance by detecting anomalies earlier and with greater sensitivity. For instance, Py-PINet begins to flag anomalies as early as day 85, well before the peak anomaly occurs around day 95. This early detection provides critical pre-warning, allowing for timely intervention to mitigate potential risks. Moreover, Py-PINet exhibits minimal missed detections, with all major anomalies correctly identified and clearly marked, ensuring comprehensive anomaly coverage. In contrast, Baseline suffers from notable missed detections, especially during the early stages of the anomaly (e.g., day 90–95), where several points exceeding the threshold are not flagged. This indicates a lack of sensitivity in Baseline's anomaly detection mechanism, which could result in delayed responses to critical issues. Additionally, Py-PINet maintains stability in healthy conditions, with low uncertainty and a narrower confidence interval, avoiding false positives. On the other hand, Baseline shows higher uncertainty and broader fluctuations in healthy conditions, increasing the likelihood of false alarms. Overall, Py-PINet outperforms baseline by providing earlier and more accurate anomaly detection, making it a more reliable choice for prognostics and health management.

## 4. Conclusions

This study developed an intelligent anomaly detection framework, demonstrating significant potential in enhancing the safety, reliability, and operational efficiency of offshore wind turbine drivetrain systems. Integrating a physics-informed probabilistic neural network (PINN), into the AI-driven model effectively allow physical insights to be in incorporated to address prediction uncertainties and ensure the reliability of failure alarms. The analysis results highlight the model's superior capability in detecting anomalies with high sensitivity and precision, particularly in prognosis, providing a robust pre-warning mechanism, enabling timely interventions to mitigate potential risks and improving the safety and stability of offshore wind turbines while optimizing maintenance strategies.

**Acknowledgement**

This research is part of the ULTIMATE project funded by the UKRI Marie Skłodowska-Curie Postdoctoral Fellowship [UKRI/EPSRC: EP/Y014235/1], State Key Laboratory of Mechanical System and Vibration [China, No. MSV202411] and the Shanghai Science and Technology Innovation Action Plan [China, No. No.24ZR1454800].

## References


Arefin, Sk. Shezan, and Md. Fatin Ishraque. 2023. “Wind Energy and Future Trends.” In *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. https://doi.org/10.1016/B978-0-323-93940-9.00090-6.

Barter, Garrett E., Latha Sethuraman, Pietro Bortolotti, Jonathan Keller, and David A. Torrey. 2023. “Beyond 15 MW: A Cost of Energy Perspective on the next Generation of Drivetrain Technologies for Offshore Wind Turbines.” *Applied Energy* 344 (August). https://doi.org/10.1016/j.apenergy.2023.121272.

Campoverde-Vilela, Lorena, María Del Cisne Feijóo, Yolanda Vidal, José Sampietro, and Christian Tutivén. 2023. “Anomaly-Based Fault Detection in Wind Turbine Main Bearings.” *Wind Energy Science* 8 (4): 557–74. https://doi.org/10.5194/wes-8-557-2023.

Mehlan, Felix C., and Amir R. Nejad. 2023. “Rotor Imbalance Detection and Diagnosis in Floating Wind Turbines by Means of Drivetrain Condition Monitoring.” *Renewable Energy* 212 (August):70–81. https://doi.org/10.1016/j.renene.2023.04.102.

Moghadam, Farid K., and Amir R. Nejad. 2022. “Online Condition Monitoring of Floating Wind Turbines Drivetrain by Means of Digital Twin.” *Mechanical Systems and Signal Processing* 162 (January). https://doi.org/10.1016/j.ymssp.2021.108087.

Shanbr, Suliman, Faris Elasha, Mohamed Elforjani, and Joao Teixeira. 2018. “Detection of Natural Crack in Wind Turbine Gearbox.” *Renewable Energy* 118 (April):172–79. https://doi.org/10.1016/j.renene.2017.10.104.

Xiang, Ling, Penghe Wang, Xin Yang, Aijun Hu, and Hao Su. 2021. “Fault Detection of Wind Turbine Based on SCADA Data Analysis Using CNN and LSTM with Attention Mechanism.” *Measurement: Journal of the International Measurement Confederation* 175 (April). https://doi.org/10.1016/j.measurement.2021.109094.

Xu, Z., M. Bashir, W. Zhang, Y. Yang, X. Wang, and C. Li. 2022. “An Intelligent Fault Diagnosis for Machine Maintenance Using Weighted Soft-Voting Rule Based Multi-Attention Module with Multi-Scale Information Fusion.” *Information Fusion* 86–87. https://doi.org/10.1016/j.inffus.2022.06.005.

Xu, Zifei, Kaicheng Zhao, Jin Wang, and Musa Bashir. 2024. “Physics-Informed Probabilistic Deep Network with Interpretable Mechanism for Trustworthy Mechanical Fault Diagnosis.” *Advanced Engineering Informatics* 62 (October). https://doi.org/10.1016/j.aei.2024.102806.

Yan, Junshuai, Yongqian Liu, and Xiaoying Ren. 2023. “An Early Fault Detection Method for Wind Turbine Main Bearings Based on Self-Attention GRU Network and Binary Segmentation Changepoint Detection Algorithm.” *Energies* 16 (10). https://doi.org/10.3390/en16104123.

Zhang, Dahai, Liyang Qian, Baijin Mao, Can Huang, Bin Huang, and Yulin Si. 2018. “A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGboost.” *IEEE Access* 6 (March):21020–31. https://doi.org/10.1109/ACCESS.2018.2818678.

Zhang, Zhen You, and Ke Sheng Wang. 2014. “Wind Turbine Fault Detection Based on SCADA Data Analysis Using ANN.” *Advances in Manufacturing* 2 (1): 70–78. https://doi.org/10.1007/s40436-014-0061-6.

Zhou, Haoyi, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. 2020. “Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting,” December. http://arxiv.org/abs/2012.07436.